MULTILEVEL MODELS: AN APPLICATION TO THE BETA-CONVERGENCE MODEL

Coro CHASCO[*] and Ana Mª LÓPEZ[*]

Abstract – *Many kinds of data in the social sciences have a hierarchical, multilevel or clustered structure. For example, municipalities are grouped into regions; regions are formed within countries; and quite often, countries belong to supra-national organizations. Once groupings are established, they will tend to become differentiated, and this differentiation implies that the group and its members both influence and are influenced by group membership. To ignore this relationship is to risk overlooking the importance of group effects and it may also render invalid many of the traditional statistical analysis techniques used for studying data relationships. In this paper, we specify a basic two-level model for a conditional beta-convergence model of a sample of European NUTS-2 regions. Specifically, we test for the role of regional decentralization (country-level variable) on regional income growth, since it has been suggested that countries with a governmental form of regional decentralization foster innovation and economic growth.*

Keywords**:** MULTILEVEL MODELS, HIERARCHICAL MODELS, CONVERGENCE, EUROPEAN REGIONS, DECENTRALIZATION, SPATIAL EFFECTS.

JEL Classification**:** C21, C23, O18, R15.

Acknowledgements: This work has been carried out with the financial support of the Instituto L.R. Klein-CentroGauss (Universidad Autónoma de Madrid, Spain) and the Project CCG08-UAM/HUM-4173 (Consejería de Educación, Comunidad de Madrid). Coro Chasco also acknowledges financial support from the Spanish Ministry of Education and Science SEJ2006-02328/ECON.

[*] Department of Applied Economics, Universidad Autónoma de Madrid (Spain).
E-mails: coro.chasco@uam.es, ana.lopez@uam.es

1. INTRODUCTION

Many kinds of data in the social sciences have a *hierarchical, multilevel* or *clustered* structure. We refer to a hierarchy as consisting of *units* grouped at different *levels*. Units at one level are recognized as being grouped, or nested, within units at the next higher level. For example, municipalities are grouped into regions; regions are formed within countries; and quite often, countries belong to supra-national organizations.

The existence of such data hierarchies is neither accidental nor ignorable. Once groupings are established, even if their establishment is effectively random, they will tend to become differentiated, and this differentiation implies that the group and its members both influence and are influenced by group membership. To ignore this relationship is to risk overlooking the importance of group effects, and may also render invalid many of the traditional statistical analysis techniques used for studying data relationships. Moreover, if the higher-level units such as countries are left out of the model, then we cannot explore potentially important questions about their effects, which we refer to as 'context'. This is the case of growth and convergence models, which should take into account not only regional factors but also national effects; e.g. economic policies, legislation, institutions or even religion (Barro and McCleary, 2003; Bräuninger and Niebuhr, 2005; Cheshire and Magrini, 2005).

Researchers have long recognized this issue in social (mainly education), medical and biological sciences through multilevel modelling (Goldstein, 2003: Raudenbush and Bryk, 2002). However, it has been largely ignored in regional science. A multilevel model is a class of variance component model that takes hierarchical data structure into account and that makes it possible to incorporate variables from all levels (or spatial scales). In this paper, we present the two-level model, which is the basic version of the multilevel model. In fact, it is a kind of random-effects variance components model in which the disturbance has a group component (for a spatially aggregated level) and an individual component (for a spatially disaggregated level).

We also illustrate its performance estimating a β-convergence model in order to explain income growth in the EU regions during the period 1992–2006. We test for the importance of national effects jointly with regional variables. Specifically, we test for the impact of regional decentralization (country-level variable) on regional income growth. It has been suggested that countries with regional decentralization foster innovation and economic growth (Darbi et al., 2003) because certain decentralized services, such as education or health care, are growth-enhancing. As long as these types of public expenditure are often provided by regional governments, promoting equalization, they also involve consequences for economic growth and efficiency (Currais and Rivera, 1999; Barro and Sala-i-Martin, 1999). Regional decentralization should lead to fiscal equalization within countries regarding the more efficient provision of public services by sub-national governments. In addition, we also consider spatial effects to test for the hypothesis of convergence clubs in the EU.

The paper is organized as follows. Section 2 contains a description of the specification and estimation of the two-level model. In Section 3, we present the empirical results of a two-level conditional beta-convergence model of regional income growth, in which a country level variable (regional decentralization) is included. The conclusions are set out in Section 4.

2. THE MULTILEVEL MODEL FOR SPATIAL DATA

2.1. Single-level relationships

We begin by presenting several models describing a single-level relationship. In spatial analysis, we refer to 'level' as a synonym of 'spatial scale'. First, we specify a single-level model at a spatially aggregated level (j), which will be considered hereafter as the country-level:

$$\begin{aligned} &y_j = \beta_0 + \beta_1 x_j + \varepsilon_j \text{ ; for } j = 1, 2, ..., J \text{ (countries)} \\ &\varepsilon_j \sim N\left(0, \sigma_\varepsilon^2\right) \end{aligned} \tag{1}$$

where standard interpretations can be given to the intercept (β_0), slope (β_1) and residual (ε).

Conducting the analysis on an aggregated spatial scale (e.g the country-level) discards all the within-group information proportioned by disaggregated spatial scale variables (e.g. region-level), which may be as much as 80% or 90% of the total variation. As a consequence, relations between spatially aggregated variables are often much stronger and they may be very different from the relation between the disaggregated variables. This is called the Modifiable Areal Unit Problem (or MAUP; see Arbia 1989).

So as not to lose valuable information, we can disaggregate all higher order variables to a lower spatial scale (e.g. regions) in order to specify a model at the regional level only (i):

$$\begin{aligned} &y_i = \beta_0 + \beta_1 x_i + \varepsilon_i \text{ ; for } i = 1, 2, ..., n \text{ (regions)} \\ &\varepsilon_i \sim N\left(0, \sigma_\varepsilon^2\right) \end{aligned} \tag{2}$$

In this case, it is common for regions in different countries to be more or less independent while regions in the same country will be closer or more similar, sharing values on many more variables. Thus, we cannot use the assumption of independence of observations. Moreover, non-observed variables will vanish into the error term causing within-group correlation between disturbances.

If we explicitly want to consider the common features shared by regions in one and the same country, we can estimate panel data models, either fixed-effects or random-effects. The following specification simultaneously describes the relationships for several countries and their corresponding regions:

$$y_{ij} = \beta_{0j} + \beta_{1j} x_{ij} + \varepsilon_{ij}$$
$$\text{for } i = 1,2,...,n \text{ (regions) ; } j = 1,2,...,J \text{ (countries)} \tag{3}$$
$$\varepsilon_{ij} \sim N\left(0, \sigma_{\varepsilon}^{2}\right)$$

This is a kind of panel data model where j refers to the aggregated spatial level 2 units (countries) and i to the (disaggregated) spatial level 1 units (regions). As it stands, it is still a single level model, albeit describing a separate relationship for each country.

In some situations, for example where there are few countries and where interest centres on just those countries in the sample, we may analyse this model by fitting all the ($2n$+1) parameters as a fixed-effects model, assuming a common 'within-regions' residual variance and separate lines for each country. The *fixed-effects model* has some potential drawbacks. First, if the sample sizes within groups are small, the estimates of the group effects may be unreliable. Second, if there are J groups to be compared, then (J–1) parameters are required to capture group effects. Third, if J is large, this entails estimating a large number of parameters.

The fixed-effects model does not allow us to make inferences beyond the groups in our sample; nor can it estimate the effects of country-level variables separately since they are conflated with the general country effects. For example, if (J – 1) dummy variables are present in the previous regression model in order to estimate country fixed-effects (β_{0j}), we cannot additionally estimate the effect of country-level characteristics – such as regional decentralization – on income growth rates. This is because any country-level variable can be expressed as a linear combination of the (J – 1) dummies.

The random-effects model estimates group (country) effects as random terms. This model allows us to focus not just on the countries included in the sample, but on a wider 'population' of countries. We then need to regard the chosen countries as giving us information about the characteristics of all the countries in the population. In particular, such a sample can provide estimates of the variation and covariation between countries in the slope and intercept parameters and will allow us to compare countries with different characteristics. Whether the levels are fixed or random depends on how these levels are chosen in any given experiment.

Both fixed-effects and random-effects panel data models are estimated at the individual level (regions) and they cannot incorporate higher level variables (countries) in the same specification. Only multilevel models take hierarchical structures into account and make it possible to incorporate variables from all levels (or spatial scales). The hierarchical or multilevel model is a kind of random-effects model and it also belongs to the wider family of random-coefficients or variance components models. It assumes that the dataset being analysed consists of a hierarchy of different populations whose differences relate to that hierarchy. Thus it also assumes that the slope of each individual (region) variable depends linearly on the class (country) variable.

2.2. Two-level model

To make (3) into a genuine two-level model we first consider it as a *random-effects model* by letting β_{0j}, β_{1j} become random variables, usually following a Normal distribution:

$$\begin{cases} \beta_{0j} = \beta_0 + u_{0j} \\ \beta_{1j} = \beta_1 + u_{1j} \end{cases} \tag{4}$$

where β_0, β_1 are fixed coefficients and u_{0j}, u_{1j} are random variables:

Then, expression (3) becomes in a *random coefficient regression model:*

$$\begin{aligned} & y_{ij} = \beta_0 + \beta_1 x_{ij} + \left(u_{0j} + u_{1j} x_{ij} + \varepsilon_{0ij}\right) \\ & \begin{cases} u_j \sim N\left(0, \sigma_u^2\right) \\ \varepsilon_{0ij} \sim N\left(0, \sigma_{\varepsilon 0}^2\right) \end{cases} \end{aligned} \tag{5}$$

We shall require an extra suffix (the '0') in the level 1 residual term (ε_{0ij}) to represent the random variation at level 1 (e.g. in a three-level model the variance structure would be more complex and the regression coefficients could also vary across level 2 units). The random variables (ε_{0ij}, u_{0j}, u_{1j}) are referred to as 'residuals' or 'errors'; in the case of a single level model, the level 1 residual ε_{0ij} becomes the usual linear model error term.

That is to say, model (5) is also a *variance component model* in which the disturbance has group-country components (u_{0ij}, u_{1ij}) and an individual-region component (ε_{0ij}). The region error terms are assumed to be independent across regions (i, t):

$$\varepsilon_{0ij} \begin{cases} E\left(\varepsilon_{0ij}\right) = 0 \\ \operatorname{var}\left(\varepsilon_{0ij}\right) = \sigma_{\varepsilon 0}^2 \qquad \forall i, t = 1, 2, ..., n \text{ (regions)} \\ \operatorname{cov}\left(\varepsilon_{0ij}, \varepsilon_{0tj}\right) = 0 \end{cases} \tag{6}$$

The country error term components are perfectly correlated within countries (σ_{u0}^2, σ_{u1}^2) but independent between countries (j, s). Some countries might be more homogenous than others; e.g. when they are grouped in clusters. This means that the variance of the country components could differ. In addition, the intercept and slope error term components are also correlated (σ_{u01}), though this relationship is constant across countries.

Formally:

$$u_{0j}, u_{1j} \begin{cases} E\left(u_{0j}\right) = E\left(u_{1j}\right) = 0 \\ \operatorname{var}\left(u_{0j}\right) = \sigma_{u0}^2 \;\; ; \;\; \operatorname{var}\left(u_{1j}\right) = \sigma_{u1}^2 \\ \operatorname{cov}\left(u_{0j}, u_{0s}\right) = 0 \;\; ; \;\; \operatorname{cov}\left(u_{1j}, u_{1s}\right) = 0 \\ \operatorname{cov}\left(u_{0j}, u_{1j}\right) = \sigma_{u01} \end{cases} \quad ; \; \forall j, s = 1, 2, ..., J \text{ (countries)} \qquad (7)$$

The feature of this specification, which distinguishes it from standard linear models of the regression or analysis of variance type, is the presence of *more than one error term* and this implies that special procedures are required to obtain satisfactory parameter estimates. It is the structure of the random part of the model that is the key factor. In the fixed part, the variables can be measured at any level; for example in the convergence model we can measure characteristics of not only regions but also countries.

Note that each level (e.g. country) is treated as a categorical variable with different random effects. This is because our observations (level 1 in the model) are nested within these categories; for example regions are nested within countries. There are several reasons for fitting a categorical variable as a random term rather than as fixed effects:

1- When our primary interest is in the variability across the various categories rather than inferences about any single category. For example, if we want to calculate how much of the variability in income growth is due to country features and how much is residual variation due to regional differences.

2- When we have only a small sample of regions for each country, given that the random effects produced will be more conservative than the category effects produced by a fixed-effect model.

3- When the number of countries is large because the fixed-effects model must fit one dummy variable for each variable. The multilevel model is more parsimonious because we do not fit an error component (u_j) for each country but instead we estimate σ_u^2 directly. Thus, if we have 20 countries we have reduced the number of parameters needed to model the between-country variation from 20 to 1. The parsimony of multilevel models allows country level explanatory variables to be fitted jointly with the between-country variance.

4- When we need complex specifications of the variance structure of the model like more levels (spatial scales) or spatial autocorrelation effects.

2.3. Covariance matrix of the error terms of a two-level model

Equation (5) requires the estimation of two fixed coefficients (β_0, β_1) and four other parameters (variances and covariances), which are the so-called random parameters:

σ^2_{u0}, σ^2_{u1}, σ^2_{u01} and $\sigma^2_{\varepsilon 0}$

For the sake of simplicity we will just consider the intercept as a random coefficient being the slope constant; i.e. $\sigma^2_{u1} = \sigma^2_{u01} = 0$. In this case, only $\sigma^2_{u0}, \sigma^2_{\varepsilon 0} \neq 0$ and ø equation (5) can be expressed as:

$$y_{ij} = \beta_0 + \beta_1 x_{ij} + \left(u_{0j} + \varepsilon_{0ij}\right)$$
$$\begin{cases} u_{0j} \sim N\left(0, \sigma^2_u\right) \\ \varepsilon_{0ij} \sim N\left(0, \sigma^2_{\varepsilon 0}\right) \end{cases} \tag{8}$$

This is termed a variance components model because the variance of the endogenous variable, about the fixed component, is:

$$\text{var}\left(y_{ij} \middle| \beta_0, \beta_1, x_{ij}\right) = \text{var}\left(u_{0j} + \varepsilon_{0ij}\right) = \sigma^2_{u0} + \sigma^2_{\varepsilon 0} \tag{9}$$

That is, the sum of level 1 and level 2 variance. This model implies that the total variance for each region ($\sigma^2_{\varepsilon 0}$) is constant (homoskedasticity) and that the covariance of the errors between two regions (denoted by ε_{i1}, ε_{i2}) in the same country (j) is given by:

$$\text{cov}\left(u_{0j} + \varepsilon_{0i1,j}, u_{0j} + \varepsilon_{0i2,j}\right) = \text{var}\left(u_{0j}\right) = \sigma^2_{u0} \tag{10}$$

since the level 1 residuals are assumed to be independent. So the covariance of the errors between two regions in the same country j coincides with the constant inter-country variance of the errors (σ^2_{u0}).

The correlation between the error terms of two such regions (i_1, i_2) included in the same country j is therefore:

$$\rho = \frac{\text{cov}\left(u_{0j} + \varepsilon_{0i1,j}, u_{0j} + \varepsilon_{0i2,j}\right)}{\text{var}\left(u_0 + \varepsilon_{0ij}\right)} = \frac{\sigma^2_{u0}}{\sigma^2_{u0} + \sigma^2_{\varepsilon 0}} \tag{11}$$

which is referred to as the *'intra-level-2-unit correlation'* (intra-class correlation); in this case, the intra-country correlation. This correlation measures the proportion of the total variance which is between-countries.[1] The existence of a non-zero intra-unit correlation, resulting from the presence of more than one error term in the model, means that traditional estimation procedures such as ‘ordinary least squares’ (OLS), which are used for example in multiple regression, are inapplicable.

We now look in more detail at the covariance structure of the errors of a group of n_j regions in a single country. It is typified by:

$$V_{2j} = \begin{bmatrix} \sigma_{u0}^2 + \sigma_{\varepsilon 0}^2 & \sigma_{u0}^2 & .. & \sigma_{u0}^2 \\ \sigma_{u0}^2 & \sigma_{u0}^2 + \sigma_{\varepsilon 0}^2 & .. & \sigma_{u0}^2 \\ \dots\dots\dots & \dots\dots\dots & .. & \dots\dots\dots \\ \sigma_{u0}^2 & \sigma_{u0}^2 & .. & \sigma_{u0}^2 + \sigma_{\varepsilon 0}^2 \end{bmatrix}_{n_j} \tag{12}$$

V_{2j} is the $(n_j \times n_j)$ covariance matrix for the income growth of n_j regions in a single country j for a two-level variance component model. The subscript 2 for V indicates that it is a two-level model. As can be seen, there is homoskedasticity but intra-country autocorrelation.

For example, when considering two countries, one of three regions and one with two NUTS-2 regions, the overall covariance matrix is a block-diagonal matrix V_2. It is the covariance matrix for the endogenous variable vector y of a two-level variance components model with two level 2 units:

$$V_2 = \left[\begin{array}{ccc|cc} \sigma_{u0}^2 + \sigma_{\varepsilon 0}^2 & \sigma_{u0}^2 & \sigma_{u0}^2 & 0 & 0 \\ \sigma_{u0}^2 & \sigma_{u0}^2 + \sigma_{\varepsilon 0}^2 & \sigma_{u0}^2 & 0 & 0 \\ \sigma_{u0}^2 & \sigma_{u0}^2 & \sigma_{u0}^2 + \sigma_{\varepsilon 0}^2 & 0 & 0 \\ \hline 0 & 0 & 0 & \sigma_{u0}^2 + \sigma_{\varepsilon 0}^2 & \sigma_{u0}^2 \\ 0 & 0 & 0 & \sigma_{u0}^2 & \sigma_{u0}^2 + \sigma_{\varepsilon 0}^2 \end{array}\right] \tag{13}$$

This ‘block-diagonal’ structure reflects the fact that the covariance between regions in different countries is zero, and clearly extends to any number of level 2 units. A more compact way of presenting this matrix, which we shall use again is:

$$V_2 = \left[\begin{array}{c|c} \sigma_{u0}^2 J_3 + \sigma_{\varepsilon 0}^2 I_3 & 0 \\ \hline 0 & \sigma_{u0}^2 J_2 + \sigma_{\varepsilon 0}^2 I_2 \end{array}\right] \tag{14}$$

where I_n is the $(n \times n)$ identity matrix and J_n is the $(n \times n)$ matrix of ones.

[1] In a model with three levels, say with countries, regions and cities, we will have two such correlations: the intra-country correlation measuring the proportion of variance that is between-countries and the intra-region correlation measuring that between regions.

In general, for a total number of J countries and n_j elements, the covariance matrix of a two-level variance component model is:

$$V_2 = \left[\begin{array}{c|c|c|c} \sigma_{u0}^2 J_{n_1} + \sigma_{\varepsilon 0}^2 I_{n_1} & 0 & .. & 0 \\ \hline 0 & \sigma_{u0}^2 J_{n_2} + \sigma_{\varepsilon 0}^2 I_{n_2} & .. & 0 \\ \hline ... & ... & .. & ... \\ \hline 0 & 0 & .. & \sigma_{u0}^2 J_{n_J} + \sigma_{\varepsilon 0}^2 I_{n_J} \end{array}\right] \quad (15)$$

In single-level OLS models, σ_{u0}^2 is zero and this covariance matrix then reduces to the standard form, $\sigma_{\varepsilon 0}^2 I_n$, where $\sigma_{\varepsilon 0}^2$ is the constant (single) error variance.

In Goldstein (2003, pp. 18–19), the covariance matrix of the error terms of the general two-level model is derived. This model is an extension of model (5) that considers random intercept and slope and also incorporates new fixed explanatory variables. A complex covariance matrix of the error terms arises, including two other covariance matrices for the random coefficients (one for the country-level and the other for the region-level).

2.4. Parameter estimation methods for multilevel models

The *Iterative Generalized Least Squares (IGLS)* is a widely-used estimation method for multilevel models. It is a two-stage process for estimating the fixed and random parameters (the variances and covariances of the random coefficients) in successive iterations (Goldstein 2003). The iteration process is what mainly distinguishes this process from the typical GLS estimation method for the random-effects panel data models.

We consider the simple two-level variance components model presented in equation (8). Suppose that we knew the values of the variances ($\sigma_{u0}^2, \sigma_{\varepsilon 0}^2$), and so could immediately construct the block-diagonal matrix V_2, which we will refer to simply as V. We can then immediately apply the usual Generalized Least Squares (GLS) estimation procedure to obtain the estimator for the fixed coefficients (β_0, β_1):

$$\tilde{\beta}^{mcg} = \left(X'V^{-1}X\right)^{-1} X'V^{-1}y$$

where, in this case: (16)

$$X = \begin{bmatrix} 1 & x_{11} \\ 1 & x_{21} \\ .. & .... \\ 1 & x_{n_j J} \end{bmatrix} \; ; \; y = \begin{bmatrix} y_{11} \\ y_{12} \\ ...... \\ y_{n_j J} \end{bmatrix}$$

with J level 2 units (countries), n_j level 1 units (regions) in the j-th level 2 unit (country) and $n_j J = n$. When the residuals have Normal distributions, GLS estimates also yield maximum likelihood (ML) estimates.

Our estimation procedure is *iterative in order to reach feasible estimates* for the generally unknown parameters. This process takes place in the following steps:

1- We would usually start from 'reasonable' estimates of the fixed parameters. Typically these will be those from an initial OLS fit (that is, assuming $\sigma_{u0}^2 = 0$) to give the OLS estimates of the fixed coefficients ($\hat{\beta}^{ols}$). From these we form the *'raw' residuals:*

$$e_{ij} = y_{ij} - \left(\hat{\beta}_0^{ols} - \hat{\beta}_1^{ols} x_{ij} \right) \tag{17}$$

The vector of raw residuals is written as:

$$e = \left\{ e_{ij} \right\} \tag{18}$$

2- If we form the cross-product matrix $e \cdot e'$ we see that the expected value of this is simply V: E($e \cdot e'$) = V.

3- We can rearrange this cross product matrix as a vector by stacking the columns one on top of the other, which is written as $vec(e \cdot e')$ and similarly we can construct the vector $vec(V)$ for the error covariance matrix. For the example given in expression (8), either $vec(e \cdot e')$ or $vec(V)$ have $3^2 + 2^2 = 13$ elements:

$$vec\left(e \cdot e'\right) = \begin{bmatrix} e_{11}^2 \\ e_{21}e_{11} \\ e_{31}e_{11} \\ \hdashline e_{11}e_{21} \\ e_{21}^2 \\ e_{31}e_{21} \\ \hdashline e_{11}e_{31} \\ e_{23}e_{31} \\ e_{31}^2 \\ \hdashline e_{12}^2 \\ e_{12}e_{22} \\ \hdashline e_{22}e_{12} \\ e_{22}^2 \end{bmatrix} \; ; \; vec\left(V\right) = \begin{bmatrix} \sigma_{u0}^2 + \sigma_{\varepsilon 0}^2 \\ \sigma_{u0}^2 \\ \sigma_{u0}^2 \\ \hdashline \sigma_{u0}^2 \\ \sigma_{u0}^2 + \sigma_{\varepsilon 0}^2 \\ \sigma_{u0}^2 \\ \hdashline \sigma_{u0}^2 \\ \sigma_{u0}^2 \\ \sigma_{u0}^2 + \sigma_{\varepsilon 0}^2 \\ \hdashline \sigma_{u0}^2 + \sigma_{\varepsilon 0}^2 \\ \sigma_{u0}^2 \\ \hdashline \sigma_{u0}^2 \\ \sigma_{u0}^2 + \sigma_{\varepsilon 0}^2 \end{bmatrix} \tag{19}$$

4- The relationship between these vectors can be expressed as the following linear model:

$$vec\left(e \cdot e'\right) = vec\left(V\right) + R$$

$$\begin{bmatrix} e_{11}^2 \\ e_{21}e_{11} \\ \cdots\cdots \\ e_{22}^2 \end{bmatrix} = \begin{bmatrix} \sigma_{u0}^2 + \sigma_{\varepsilon 0}^2 \\ \sigma_{u0}^2 \\ \cdots\cdots\cdots\cdots \\ \sigma_{u0}^2 + \sigma_{\varepsilon 0}^2 \end{bmatrix} + R = \sigma_{u0}^2 \begin{bmatrix} 1 \\ 1 \\ .. \\ 1 \end{bmatrix} + \sigma_{\varepsilon 0}^2 \begin{bmatrix} 1 \\ 0 \\ .. \\ 1 \end{bmatrix} + R \tag{20}$$

where R is a residual vector. The left hand side of (18) is the vector of the raw residuals in the OLS estimation of the linear model and the right hand side contains two explanatory variables, with coefficients σ_{u0}^2, $\sigma_{\varepsilon 0}^2$, which are to be estimated.

5- The estimation involves an application of GLS using the estimated covariance matrix of $vec(e' \cdot e)$, assuming Normality, namely $2(V^{1} \otimes V^{1})$ where $\otimes$ is the Kronecker product. The Normality assumption allows us to express this covariance matrix as a function of the estimated random parameters.

6- If the Normality assumption fails to hold,[2] the resulting IGLS estimates are still consistent although not fully efficient, but the standard errors (estimated using the Normality assumption) and confidence intervals will generally not be consistent.

7- With the estimates obtained from applying GLS to expression (20) we return to expression (16) to obtain new GLS estimates of the fixed effects and so alternate between the random and fixed parameter estimation until the procedure converges, that is the estimates for all the parameters do not change from one iteration to the next.

Essentially the same procedure can be used for more complicated models (Rasbash et al., 2000).

The maximum likelihood (ML) procedure produces biased estimates of the random parameters because it takes no account of the sampling variation of the fixed parameters. This may be important in small samples and we can produce unbiased estimates by using a modification known as restricted maximum likelihood (REML). As stated in Goldstein (2003), the IGLS algorithm is readily modified to produce restricted estimates that are called

[2] For certain variance component models alternative distributional assumptions have been studied, especially for discrete response models of the kind discussed in Clayton and Kaldor (1987) and maximum likelihood estimates obtained. For more general models, however, with several random coefficients, the assumption of multivariate Normality is a flexible one, which allows a convenient parameterization for complex covariance structures at several levels.

RIGLS. This is the method we use in the paper since we work with a relatively small sample.

There are many other algorithms and estimation methods for multilevel models. This is the case of the Expectation Maximization (EM) algorithm or variants of it (Raundenbush and Bryck, 2002) or Longford (1987)'s 'Fisher scoring' algorithm. A rather different approach is to view the multilevel linear model as a Bayesian linear model (Lindley and Smith, 1972). An alternative to the full Bayes estimation, known as 'Empirical Bayes' (EB), ignores the prior distributions of the random parameters, treating them as known for purposes of inference. When Normality is assumed, these estimates are the same as IGLS or RIGLS. More recently, the full Bayesian treatment has become computationally feasible with the development of 'Markov Chain Monte Carlo' (MCMC) methods, especially Gibbs Sampling (Zeger and Karim, 1991). This has the advantage, in small samples, of taking account of the uncertainty associated with the estimates of the random parameters and can provide exact measures of uncertainty. The maximum likelihood methods tend to overestimate precision because they ignore this uncertainty.

3. MODEL AND DATA

3.1. Specification of the β-convergence model

Our aim is to determine to what extent decentralized countries in Europe (from a political and economic point of view) fostered more economic growth in their corresponding regions than countries with a classical unitary state. For this purpose, we specify an income growth model of the EU regions taking into account the neo-classical growth model (Solow, 1956), which is considered as a natural starting point for the analysis of regional disparities, especially in Europe (Fingleton, 2003). The neo-classical model predicts that the growth rate of a region is positively related to the distance that separates it from its steady state. That is to say, if all regional economies are structurally identical and have access to the same technology, they are characterized by the same steady state, and differ only by their initial conditions.

In the case of the European regions, as their initial conditions are not similar, we initially propose a *conditional β-convergence model,* in which per capita income growth in period (0, T) is a function of per capita income at time 0 and two control variables that proxy the differences in steady-state positions across different economies, employment (E_i) and regional decentralization (D_i). We also control for the presence of spatial effects (S_i), mainly in the form of spatial heterogeneity (spatial clubs convergence hypothesis). In general terms, the formal expression is:

$$
\begin{aligned}
&G_i = f(g91_i, E_i, D_i, S_i, \varepsilon_i) \\
&\varepsilon_i \sim N(0, \sigma_\varepsilon^2)
\end{aligned}
\qquad (21)
$$

where G_i is the endogenous variable, which is measured as the average of the natural log of per capita GVA growth rate in region *i* during the period 1991–2006. In the present work, data are extracted from the Cambridge Econometrics regional database,[3] which provides comparable regional data at NUTS-2 (Nomenclature of Territorial Units for Statistics, established by Eurostat) level on level on real gross value-added (GVA).[4] The choice of the NUTS-2 level is the appropriate one when analysing regional decentralization in countries, since it is the spatial unit with decision capacity. We have selected a sample of 233 NUTS 3 regions in 20 EU countries.[5] The sample includes information about Austria (9 units), Belgium (11 units), the Czech Republic (8 units), Denmark (5 units), Finland (4 units), France (21 units), Germany (39 units), Greece (9 units), the Netherlands (12 units), Hungary (7 units), Ireland (2 units), Italy (19 units), Luxembourg (1 unit), Poland (16 units), Portugal (4 units), Slovakia (4 units), Slovenia (2 units), Spain (15 units), Sweden (8 units) and the United Kingdom (37 units).

In the group of explanatory variables, $g91_i$ represents the initial conditions, which is proxied by per capita GVA in NUTS-2 region *i* in the year 1991 (in natural logarithms). Following the neoclassical paradigm, we expect that less developed regions are catching up with the richer ones, so that regional levels of per capita income tend to converge over the long run because of diminishing returns on capital. In a competitive environment, regional labour and capital mobility, as well as regional trade, will also work in favour of factor price convergence, reinforcing the negative relation between growth and regional inequality. However, other schools of thought tend to agree with Myrdal's basic claim (1957) that growth is a spatially cumulative process, which is likely to increase inequalities (divergence). Therefore, a negative value for the slope coefficient β indicates convergence of per capita GVA across regions within a given time period, while a positive value indicates divergence. In the group of the selected NUTS-2 regions, a negative sign for β is expected, as most of the regions with lower per capita income growth in the 1991–2006 period were those with higher per capita GDP in 1991, and vice versa.

The hypothesis of conditional convergence is tested controlling for permanent cross-regions differences that could potentially explain regional growth income. These variables, which are usually referred to the first moment of time, allow us to proxy the differences in steady-state positions across different economies. These control variables represent all other effects that contribute or weaken economic growth (Mella and Chasco, 2006). Occupation rate (*E*) is one of the several control variables used in the literature. It is measured as the percentage of population in employment in 1991. We expect a

[3] See in Table 1 a complete specification of all the variables.
[4] GVA equals GDP net of taxes on and subsidies for production.
[5] From the total group of 253 NUTS-3 regions in the EU-25, we have omitted some units with missing data and the 'islands'; i.e. those regions with no spatially contiguous neighbours. Some regions with extremely high values of income growth rates have also been excluded. This is why Cyprus, Malta, Estonia, Latvia and Lithuania are not present leading to a regional sample of 233 NUTS-2 in 20 EU countries.

positive sign for its estimated coefficient due to the fact that income growth is the final result of an accumulation process of human resources in a region.

Table 1: Definition of the variables

	Variable	***Level***	***Source***
Dependent variable:			
G	Average per capita GVA growth rate, period 1991–2006 (in natural logarithms)	NUTS-2 regions	European Regional Database (Cambridge Econometrics)
Explanatory variables:			
A) Initial conditions:			
*g*91	GVA per capita in 1991 (in natural logarithms)	NUTS-2 regions	European Regional Database (Cambridge Econometrics)
B) Control variables:			
E	Percentage of employees over total population in 1991 (in natural logarithms)	NUTS-2 regions	European Regional Database (Cambridge Econometrics)
D	Forms of regional government in 1991 (1: fully centralized; 2: devolving unitary countries; 3: regionalized or federal countries)	Countries	Self-elaboration from Russel Barter (2000)
*Wg*91	Spatial lag variable of GVA per capita in 1991 (in natural logarithms)	NUTS-2 regions	Self elaboration from European Regional Database (Cambridge Econometrics)
WE	Spatial lag variable of Percentage of employees over total population in 1991 (in natural logarithms)	NUTS-2 regions	Self elaboration from European Regional Database (Cambridge Econometrics)
x	X-coordinate (East-West direction)	NUTS-2 regions	Self elaboration
y	Y-coordinate (North-South direction)	NUTS-2 regions	Self elaboration
C) Spatial split variable:			
REG	Spatial regimes (1: core regions; 0: periphery regions)	NUTS-2 regions	Self-elaboration

Regional decentralization (*D*) is also a control variable in this model. It has been defined as a categorical country-level variable since it adopts the same values for the NUTS-2 regions belonging to the same country. The categories have been defined from the typology established by Rusell Barter (2000) for the forms of regional government[6]: value 1 for those classic centralized states with no powers for the regions in the year 1991 (Eastern countries, Finland, Greece, Ireland, Luxembourg, Portugal, Sweden and UK); value 2 for devolving centralized states with limited powers for the regions in the year 1991 (France, the Netherlands and Denmark); and value 3 for regionalized or federal states with advanced and wide-ranging powers for the regions (Austria, Germany, Belgium, Spain and Italy).

Finally, we have also considered the presence of spatial effects, which invalidate the assumption of independence of the error terms in convergence models (Rey and Montouri, 1999). In fact, spatial spillovers (spatial autocorrelation and/or spatial heterogeneity) should be taken into account to avoid bias in the resulting estimates. An option when considering spatial spillovers of regional income growth is to include exogenous spatial lag as environmental explanatory variables (Le Gallo et al., 2003; Fingleton and López-Bazo, 2006). This is also the solution proposed by Morenoff (2003) for hierarchical models, since they have not yet implemented the proper estimation methods for the spatial lag and the spatial error model. Therefore, *Wg*91 and *WE* are the spatial lag variables of per capita income and employment rate (both referred to 1991).[7] Other spatial variables capable of capturing spatial trends in the data are the Earth coordinates: *x* (East - West direction) and *y* (North - South direction).

In addition, we also test for the club-convergence hypothesis (Durlauf and Johnson, 1995), which implies the existence of different regional economies (clubs) that are similar in structural characteristics and tend to converge within groups. The equilibrium that each region will reach depends on the initial conditions of the group to which they belong. The composition of the regional clubs could be defined after an exploratory spatial data analysis (ESDA) as different spatial regimes. Moran's scatterplot is a good instrument for detecting spatial clusters in a variable (Anselin, 1996). In Figure 1, we have represented the Moran's scatterplot of the per capita income in 1991 (in natural logarithms).[8] It demonstrates that in 1991, the regional income distribution appeared to be clustered in nature. That is, regions with relatively high/low income levels tended to be located near other regions with high/low income

[6] It must be said that the majority of European countries, with the exception of federal systems, are unitary in the sense that sovereignty is exclusively invested in central government. The degree to which central governments may devolve functions and power, however, clearly differs. In this paper, what we consider is decentralization of spending decisions particularly in areas such as health and education. For example, Russell Barter (2000) considers the UK as a centralized country in spite of the devolved administrations for Wales, Scotland and Northern Ireland because devolution has not so far had a major impact on the degree to which expenditure is assigned to sub-national jurisdictions.

[7] The spatial weights matrix, W, has been defined as an inverse distance matrix, such that each element w_{ij} is set equal to the inverse of the squared distance between regions i and j. Namely, region vicinity is inversely determined by the relative distance that separates them.

[8] The magnitude of the Moran's I statistic for spatial autocorrelation is high (I=0.56) and very significant at $p = 0.001$, which is well above its expected value under the null hypothesis of no spatial autocorrelation, $E[I] = -0.003$. Inference is based on the permutation approach (999 permutations). There is also high spatial autocorrelation in the distribution of employment rates with a Moran's I value of 0.24 ($p = 0.001$).

levels more often than would be expected as a result of purely random factors. The fist quadrant of the Moran's scatterplot represents those higher-income regions (above average) that are surrounded by higher-income neighbours. They represent the 'core' in terms of income, the rest of the NUTS-2 regions being the 'periphery'. Highlighting them in the map, splits the NUTS-2 regions into two spatial regimes (clubs) as has been detected in other studies (e.g. Canova, 2004; Ertur et al., 2006).

Figure 2 is a scatterplot of the log per capita income growth rate in 1991–2006 by the log per capita income in 1991. In this plot no distinction is made between the countries to which the NUTS-3 units belong. There is a general trend, with increasing income per capita in 1991 associated with decreasing income growth rate in the period 1991–2006. This result confirms the hypothesis of regional convergence in Europe. In addition, ø the narrowing of the between NUTS-3 variation in the income growth rate with increasing income per capita in 1991 is remarkable, depicting some interesting clusters in the area of maximum dispersion. There is a group of regions from the Eastern countries that starting from lower per capita income in 1991, exhibited a completely different behaviour in terms of income growth. First, most NUTS-2 from the Czech Republic and Hungary, which are centralized countries, had the lowest income growth rates (down-divergence). Second, NUTS-2 from Poland and Slovakia, which are actually devolving centralized countries, attained higher (above the average) income growth (convergence). Third, there is another cluster of NUTS-2 from Ireland and East-Germany that achieved the highest income growth with an initial per capita income above the European average (upper-divergence). Note that the outstanding clusters (first and third) are formed by NUTS-2 from countries with regional decentralization, with the exception of Ireland.

Figure 1: Moran scatterplot (left) and map (right) of regional log per capita GVA in 1991: selection of regions in the first quadrant

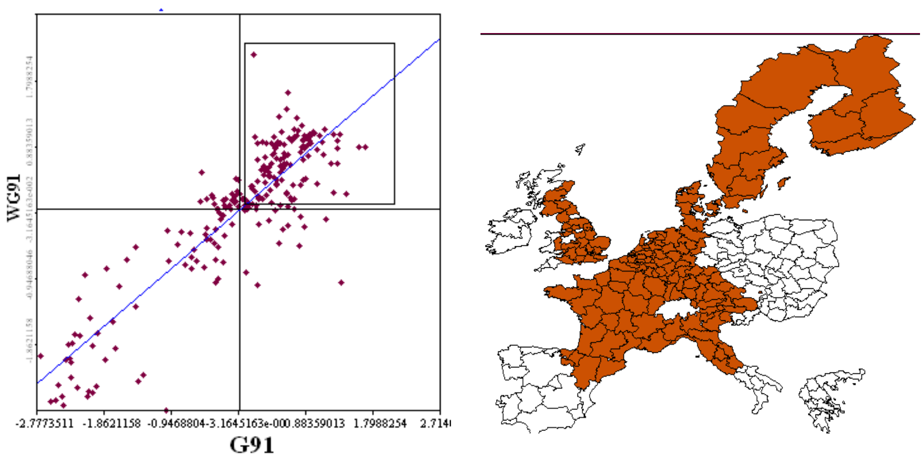
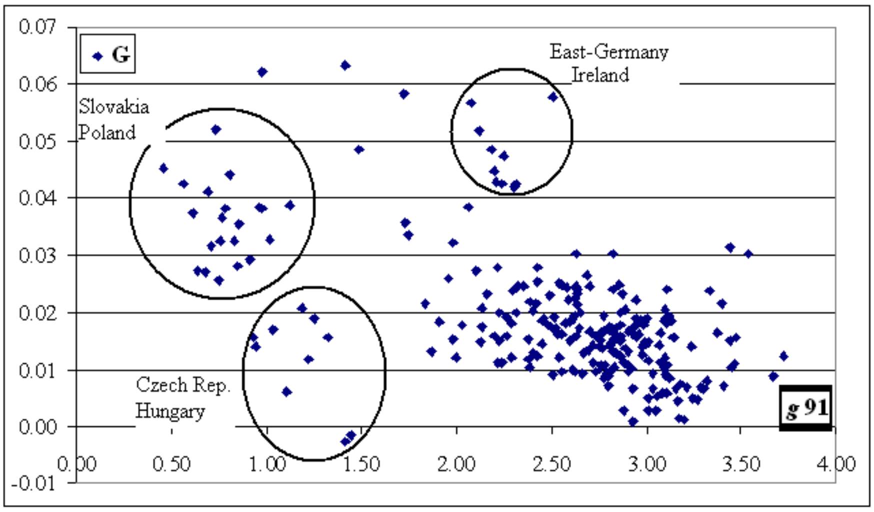


Figure 2: Scatterplot of regional income growth in 1991–2006 (D) against per capita income in 1991 (g91)



3.2. Estimation of a spatial multilevel beta-convergence model

Our initial model is a one-level panel data of the conditional β-convergence model which relates regional per capita GVA growth rate (G) with per capita GVA in 1991 ($g91$) and the following control variables: employment rate (E) and the spatial lag variable of per capita income in 1991 ($Wg91$)[9]:

$$G_{ij} = \beta_0 + \beta_1 g91_{ij} + \beta_2 E_{ij} + \beta_3 Wg91_{ij} + \varepsilon_{ij}$$
$$\text{for } i = 1,2,...,233 \text{ (NUTS-2 regions) ; } j = 1,2,...,20 \text{ (countries)} \qquad (22)$$
$$\varepsilon_{ij} \sim N\left(0, \sigma_\varepsilon^2\right)$$

In this model, no categorical country-level has been specified yet, so it can be estimated by Ordinary Least Squares (OLS).[10] As is shown in Table 2, all the coefficients are quite significant. The multicollinearity figure is 30, which is considered acceptable. The Jarque-Bera non-normality statistic on the residuals takes on a significantly high value. Consequently, we will treat with caution the results of the misspecification tests depending on the normality assumption. The Kelejian-Robinson test for spatial autocorrelation, which is not affected by non-normal errors, is not significant ($p = 0.943$). Nevertheless, both the Koenker-Basset test and the White test are quite significant showing the existence of an unspecified form of heteroskedasticity which is probably due (as shown in the ESDA) to the existence of two spatial regimes in per capita

[9] The spatial lag variable of employment rate (WE) was removed from the model since it is not statistically significant at 95%.
[10] Note that these estimates coincide with the first iteration of IGLS.

income (Figure 1). This result is confirmed by the spatial Chow test[11] on the null hypothesis where the coefficients are the same in both spatial regimes, which is also clearly rejected at 99%.

Therefore, OLS errors are not random and the sources of this non-randomness could be the existence of two spatial regimes, as suggested by ESDA, the White test and the spatial Chow test. Model (22) will also lead to bias results because it ignores country - fixed or random - effects. Since we want to test for the incidence of a country-level variable - forms of regional government (*D*) - on income growth, we next specify the following random-effects multilevel model:

$$\begin{aligned} G_{ij} &= \beta_{0j} + \beta_1 g91_{ij} + \beta_2 E_{ij} + \beta_3 Wg91_{ij} + \beta_4 WE_{ij} + \beta_5 D\_devolv_j + \\ &\quad + \beta_6 D\_regional_j + \varepsilon_{ij} \\ \beta_{0j} &= \beta_0 + u_{0j} \\ u_j &\sim N\left(0, \sigma_u^2\right) \\ \varepsilon_{ij} &\sim N\left(0, \sigma_\varepsilon^2\right) \end{aligned} \tag{23}$$

This model, which introduces a new country-level categorical variable (*D*), has been estimated by RIGLS; the estimation was completed at iteration 12. As stated before, one of the strengths of multilevel modelling is the ability to estimate correctly between-country variation and also include country-level regressors. In order to assess whether model (23) is more efficient than model (22) we compute the likelihood ratio test[12] (LR) of H0: $\sigma_{u0}^2 = 0$. That is to say, we compare the multilevel model with the basic model, where σ_{u0}^2 is constrained to equal zero. The value of the likelihood ratio statistic, obtained from the two models log-likelihoods is 48.27 (p = *0.000*). We conclude that there is significant variation between countries.

Recall that model (23) amounts to fitting a set of parallel straight lines to the data from the different countries. The slopes of the lines are all the same, and the fitted values of the common slopes are – 0.0178 (*g*91), 0.0377 (*E*), – 0.0099 (*W*1*g*91), 0.0211 (*WE*), 0.0133 (*D_devolv*) and 0.0101 (*D_regional*). As we can see in Table 2, all the estimates are clearly significant at 95% with the exception of the two dummies of regional decentralization. The intercepts for the different countries are the country-level residuals (u_{0j}) and these are distributed around zero with a variance of $\hat{\sigma}_{u0}^2 = 2.9038 \cdot 10^{-4}$ (between-country estimated variance). Obviously the actual data points do not lie exactly on the

---

[11] The spatial Wald-Chow test was proposed by Anselin (1990). It is based on an asymptotic Wald statistic distributed as a χ^2 distribution with [(m – 1).k] degrees of freedom (m being the number of regimes and k the number of estimated coefficients).

[12] This LR test corresponds to twice the difference between the log likelihood in the multilevel model and the log likelihood in a standard regression model with the same set of explanatory variables. It is distributed as a χ^2 variate with one degree of freedom.

straight lines; they vary about them by amounts given by the NUTS-2 region-level residuals (ε_{ij}) which have a variance estimated as $0.574{\cdot}10^{-4}$.

Table 2: Estimation results for different models of regional per capita GVA growth

Model	Standard model	Two-level model	Two-level model (Core regions)	Two-level model (Periphery regions)
Estimation	OLS	RIGLS	RIGLS	RIGLS
# observations	233	233	146	87
Intercept (β_0)	0.0543^{**} (0.0050)	-	-	-
Centralized states mean value (β_0)	-	0.1371^{**} (0.0131)	0.0171^{**} (0.0074)	0.0863^{**} (0.0178)
*g*91	-0.0071^{**} (0.0020)	-0.0185^{**} (0.0030)	0.0052^{**} (0.0018)	-0.0151^{**} (0.0053)
E	0.0078^{**} (0.0043)	0.0390^{**} (0.0011)	-	0.0354^{**} (0.0109)
*Wg*91	-0.0043^{*} (0.0030)	-0.0100^{**} (0.0025)	-	-
WE	-	0.0216^{**} (0.0088)	0.0099^{**} (0.0059)	-
x-coordinate (East-West direction)	-	-	-0.00022^{**} (0.00014)	-0.00065^{**} (0.00037)
D_devolv	-	0.0140 (0.0129)	-0.0074^{**} (0.0039)	-
D_regional	-	0.0107 (0.0106)	-0.0082^{**} (0.0035)	0.0067 (0.0106)
Between-country variance (σ_u^2)	-	$3.735{\cdot}10^{-4}$	$2.312{\cdot}10^{-5}$	$2.320{\cdot}10^{-4}$
Between-region variance (σ_ε^2)	$9.737{\cdot}10^{-5}$	$0.581{\cdot}10^{-4}$	$1.527{\cdot}10^{-5}$	$1.055{\cdot}10^{-4}$
LIK	745.493	769.626	589.308	255.553
AIC	–1,482.99	–1525.25	–1,166.62	–511.11
Jarque-Bera norm. test	111.78^{**}	-	-	-
Multicollinearity #	29	-	-	-
Koenker-Bassett test	31.00^{**}	-	-	-
White test	55.02^{**}	-	-	-
Kelejian-Rob. test	0.76	-	-	-
Spatial Chow test	3.35^{**}	-	-	-
Convergence speed (*b*)	0.8%	2.2%	Divergence	1.3%
Half-life (τ)	97 years	37 years		60 years

Notes: * Null-hypothesis rejection between 5–10 per cent of significance. ** Null-hypothesis rejection below 5 per cent of significance. *Standard* is a one-level model. *OLS* indicates ordinary least squares estimation. *RIGLS* indicates rectrictive iterative generalized least squares estimation. *LIK* is the log-likelihood value. *AIC* is the Akaike Information Criterion. *Jarque Bera norm. test* is the Jarque-Bera non-normality test on the residuals. *Multicollinearity #* is a test for multicollinearity in the regressors. *Koenker-Bassett test* is the Koenker-Bassett test for hetero-skedasticity robust to non-normality in the errors. *White test* is the White test for unspecified heteroskedasticity. *Kelejian-Rob.* is the Kelejian-Robinson test for spatial autocorrelation robust to non-normality in the errors. *Spatial Chow test* is the spatial Chow-Wald test on spatial instability of the coefficients in two regimes: core NUTS-2 regions (REG=1) and Periphery NUTS-2 regions (REG=0). *Convergence speed* is the convergence speed. *Half-life* is the time necessary for the group of cities to reach half of the variation, which separates them from their steady state.

The intra-country correlation measures the proportion of the total variance which is between-countries: $\hat{\rho} = \hat{\sigma}^2_{u0} \big/ (\hat{\sigma}^2_{u0} + \hat{\sigma}^2_{\varepsilon 0}) = 0.87$; i.e. almost 90% of the total variance in NUTS-2 region income growth may be attributed to differences between countries. This result also highlights the superiority of a random coefficients multilevel model and seems to confirm Cheshire and Magrini's (2005) assumption, i.e. that 'regions within the EU seem to behave like city-states, not as simply the spatial units from which a continental economy is constructed'.

The reference category for variable *D* is 'centralized states', and it has been defined as two dummy variables, *D_devolv* and *D_regional*, which take a value of 1 for those NUTS-2 regions in devolving centralized and regionalized/federal countries, respectively. The parameter β_0 is the mean for centralized countries whereas β_5 (*D_devolv*) and β_6 (*D_regional*) represent the mean difference in income growth between regions in devolving centralized or regionalized/federal states (respectively) and regions in a classic unitary country. We see that during the period 1991–2006, the EU NUTS-2 regions from devolving and regional/federal states grew more than regions from centralized states. Nevertheless, these results must be considered with caution because the estimated coefficients for variables *D_devolv* and *D_regional* are not statistically significant at all. In addition, the existence of two remarkable spatial regimes can also bias all the results of this model.

Using the estimated β_1 coefficient, the convergence process can be characterized by two additional concepts: (i) convergence speed, which can be defined as $b = -\ln(1+T\beta_1)/T$ (for T=15 years), and, (ii) half-life or the time necessary for the economies to reach half of the variation that separates them from their steady state: $\tau = -\ln(2)/\ln(1+\beta_1)$. In this model, the associated speed of convergence is 2.2 per cent, close to the 2 per cent usually found in the standard convergence literature, which indicates a similar process (the half-life is 37 years).

In order to test whether spatial heterogeneity might be biasing the previous outcome, we estimate model (23) for the two previously defined spatial regimes: core and periphery. In effect, results change significantly. In the core area, both *D*_devolv and *D*_regional are more or less similar in value, significant, and negative; i.e. regions in devolving centralized and regionalized/federalist countries experienced less income growth than regions in totally centralized countries. In the 'periphery' area, we have got a different result: *D_regional* achieves a positive significant coefficient. In addition, we find a convergence trend across the peripheral regions, though they are moving at a slower speed (1.1%). Nevertheless, in the core, NUTS-2 regions experienced clear divergence.

In short, it is not clear that country decentralization has had a positive impact on the economic development of European regions. If we consider them as a whole, we find some evidence in favour of decentralization, though it is not statistically significant. If we split the EU NUTS-2 into core-periphery regions, we find that only in the core is decentralization significant but in the diametrically opposite sense to that expected; i.e. the most dynamic regions were those located in more centralized countries (Finland, Sweden and the UK). Regarding peripheral regions, we find evidence in favour of more decentralized countries (e.g. Spain and Italy with respect to Eastern countries and Greece) but it is not statistically significant.

4. CONCLUSIONS

In this paper, we present the multilevel or hierarchical model as a good instrument for estimating group-level variables (e.g. countries) jointly with individual-level variables (e.g. regions). That is to say, it allows us to estimate between-country variation and country-level regressors in the same model. We illustrate its performance estimating a two-level conditional beta-convergence model in order to explain income growth in the EU regions during the period 1992–2006. In particular, we test for the importance of regional decentralization (country-level variable) jointly with classical regional variables, such as initial income, employment rate and spatial variables. Since decentralized provision and fiscal autonomy can promote the efficiency and accountability of sub-central governments, a positive impact of this variable is expected on income growth. Nevertheless, after controlling for some variables, results are really surprising for the EU regions. Overall, there is only weak evidence in favour of regional decentralization as an engine of income growth.

However, splitting the regions into a core-periphery framework, we find that in the core of the EU decentralization is no guarantee of higher income growth rates. On the contrary, we easily find dynamic regions in Luxembourg, the UK and Finland, which are considered as classic unitary states from the point of view of spending decisions particularly in areas such as health and education. In peripheral regions, the role of national decentralization in peripheral regions ø is positive at least but statistically insignificant. In fact, higher income rates are found in Ireland and some Eastern countries (mainly Poland and Slovakia) instead of the peripheral regions of highly regionalized countries such as Italy or Spain.

These astonishing results are not possibly questioning the efficiency of regional decentralization in theory, but the way in which it has been developed in Europe. Multiple forms of decentralization make it difficult to construct this variable. In addition, asymmetric and multi-speed reforms - although sometimes justified on historical grounds—have resulted in considerable added complexity and cannot be justified on economic grounds since it is probably unsustainable.

REFERENCES

Anselin L., 1996, "The Moran scatterplot as an ESDA tool to assess local instability in spatial association", *in* Fischer M., Scholten H., Unwin D. (eds.), *Spatial Analytical Perspectives on GIS*, Taylor and Francis, London, UK.

Anselin L., 1990, "Spatial dependence and spatial structural instability in applied regression analysis", *Journal of Regional Science*, vol. 30, pp. 185–207.

Arbia G., 1989, *Spatial Data Configuration in the Statistical Analysis of Regional Economics and Related Problems*, Kluwer Academic Publishers, Dordrecht.

Barro R.J., McCleary R.M., 2003, "Religion and economic growth", *American Sociological Review*, vol. 68-5, pp. 760-781.

Barro R.J., Sala-i-Martin X., 1999, *Economic Growth*, The MIT Press, Cambridge.

Bräuninger M., Niebuhr A., 2005, Convergence, spatial interaction and agglomeration effects in the EU, RePec papers.

Canova F., 2004, "Testing for convergence clubs in income per capita: a predictive density approach", *International Economic Review*, vol. 45, pp. 49-77.

Clayton D.G., Kaldor J., 1987, "Empirical Bayes estimates of age-standardised relative risks for the use in desease mapping", *Biometrics*, vol. 43, pp. 671-681.

Cheshire P., Magrini S., 2005, European urban growth: throwing some economic light into the black box, 45th European Regional Science Association Congress, Amsterdam.

Currais L., Rivera B., 1999, "Income variation and health expenditure: Evidence for OECD countries", *Review of Development Economics*, vol. 3, pp. 258-267.

Darbi J.., Muscatelli A., Graeme R., 2003, Fiscal decentralisation In Europe: A review of recent experience, *Department of Economics, University of Glasgow*, G12 9RT.

Durlauf S.N., Johnson P.A., 1995, "Multiple regimes and cross-country growth behaviour", *Journal of Applied Econometrics*, vol. 10, pp. 365–84.

Ertur C., Le Gallo J., Baumont C., 2006, "The European regional convergence process, 1980-1995: Do spatial regimes and spatial dependence matter?", *International Regional Science Review*, vol. 29, pp. 3-24.

Fingleton B., 2003, “Models and simulations of GDP per inhabitant across Europe’s regions: a preliminary view”, *in* Fingleton B. (ed.), *European Regional Growth*, Berlin and Heidelberg, Springer, Berlin,

Fingleton B., López-Bazo E., 2006, “Empirical growth models with spatial effects”, *Papers in Regional Science*, vol. 85, pp. 177-198.

Goldstein H., 2003, *Multilevel Statistical Models, third ed.,* Arnold, London.

Le Gallo J., Ertur C., Baumont C., 2003, “A spatial econometric analysis of convergence across European regions, 1980-1995”, *in* Fingleton B. (ed.) *European Regional Growth*, Springer, Berlin and Heidelberg

Lindley D.V., Smith A.F.M., 1972, “Bayes estimates for the linear model”, *Journal of the Royal Statistical Society, Series B*, vol. 34, pp. 1-41.

Longford N.T., 1987, “A fast scoring algorithm for maximum likelihood estimation in unbalanced mixed models with nested random effects”, *Biometrika*, vol. 74, pp. 817-827.

Mella J.M., Chasco C., 2006, “A spatial econometric analysis of urban growth and territorial dynamics: A case study on Spain”, *in* Reggiani A., Nijkamp P. (eds.), *Spatial Dynamics, Networks and Modeling*, Edward Elgar, pp. 319-360.

Morenoff J.D., 2003, “Neighborhood mechanisms and the spatial dynamics of birth weight”, *American Journal of Sociology*, vol. 108, pp. 976–1017.

Myrdal G., 1957, *Economic Theory and Underdeveloped Region*s, Methuen, London, UK.

Rey S., Montouri B., 1999, “US regional income convergence: a spatial econometric perspective”, *Regional Studies*, vol. 33, pp. 143–56.

Rabash J., Browne W., Goldstein H., 2000, *A User’s Guide to MLwIN, 2nd edition*, London Institute of Education.

Raudenbush S.W., Bryk A.S. (2002) *Hierarchical Linear Models: Applications and Data Analysis Methods*, Sage Publications, Thousand Oaks.

Russell-Barter W., 2000, *Regional government in England. A preliminary review of literature and research findings*, Department of the Environment, Transport and the Regions, London.

Solow R., 1956, “A contribution to the theory of economic growth”, *Quarterly Journal of Economics*, vol. 70, pp. 101–108.

Zeger S.L., Karim M.R., 1991, “Generalized linear models with random effects; A Gibbs sampling approach”, *Journal of the American Statistical Association*, vol. 86, pp. 79-86.

MODÈLES MULTI-NIVEAUX : UNE APPLICATION AU MODÈLE DE BETA-CONVERGENCE

Résumé *– De nombreuses données en science sociale possèdent une structure hiérarchique ou multi-niveaux. Par exemple, les communes font partie de régions, qui elles-mêmes font partie de pays. Ces derniers font en outre le plus souvent partie d'organisations supranationales. Une fois que ces différents regroupements ont été établis, ils sont différenciés. Cette différenciation implique que le groupe et ses membres influencent et sont influencés par l'appartenance au groupe. Ignorer ce type de relation revient à ignorer l'importance des effets de groupe, ce qui potentiellement invalide les analyses statistiques traditionnelles. Cet article s'attache alors à spécifier un modèle simple à deux niveaux pour un modèle de beta-convergence conditionnelle estimé sur un échantillon de régions européennes au niveau NUTS-2. Plus précisément, nous analysons le rôle de la décentralisation régionale (variable mesurée au niveau national) sur la croissance du revenu régional, afin de vérifier si les pays possédant une certaine forme de décentralisation régionale connaissent une croissance plus rapide.*